

Weak Convergence of Stationary Empirical Processes

Dragan Radulović

Department of Mathematics, Florida Atlantic University

Marten Wegkamp

Department of Mathematics & Department of Statistical Science, Cornell University

April 2017

Abstract

We offer an umbrella type result which extends the convergence of classical empirical process on the line to more general processes indexed by functions of bounded variation. This extension is not contingent on the type of dependence of the underlying sequence of random variables. As a consequence we establish the weak convergence for stationary empirical processes indexed by general classes of functions under alpha mixing conditions.

Running title: Weak convergence of stationary empirical processes

MSC2000 Subject classification: Primary 60F17 ; secondary 60G99.

Keywords and phrases: Integration by parts, bounded variation, empirical processes, stationary distributions, weak convergence.

1 Introduction

We consider the empirical process

$$\mathbb{Z}_n(g) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \{g(X_i) - \mathbb{E}g(X_i)\}, \quad g \in \mathcal{G}, \quad (1)$$

indexed by the class of functions \mathcal{G} . It is an obvious generalization of the classical process

$$\mathbb{G}_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \{1_{(-\infty, t]}(X_i) - F(t)\}, \quad (2)$$

in which case the indexing class is $\mathcal{G} = \{g(x) = 1_{(-\infty, t]}(x) : t \in \mathbb{R}\}$. If the underlying sequence $\{X_i\}$ is i.i.d., then the limiting behavior of the empirical processes $\mathbb{Z}_n(g)$ is well understood. The same is true for the bootstrap counterpart $\mathbb{Z}_n^*(g)$, based on an i.i.d. bootstrap sample $\{X_i^*\}$, see Van der Vaart & Wellner (1996), Dudley (1999). The theory of weak convergence of empirical processes based on independent sequences has yielded a wealth of statistical applications and, in particular, it was instrumental for establishing the weak convergence of numerous novel statistics. Often the limiting distributions of these statistics do not allow for closed form solution, in which case the bootstrap version of the process is utilized.

For empirical processes based on stationary sequences $\{X_i\}$, the situation is rather different. The classical process $\mathbb{G}_n(t)$ has been treated by numerous authors, who established weak convergence under sharp mixing assumptions, see, for instance, Rio (2000). However, nothing similar exists for more general processes. The only work that we could find in the literature, Andrews and Pollard (1994), treats more general indexing classes, but imposes very restrictive assumptions on the decay of α -mixing coefficients. This discrepancy, between the conditions needed for $\mathbb{Z}_n(g)$ and $\mathbb{G}_n(t)$, is due to fact that typical approach for proving the uniform limiting theorems heavily relies on the estimation of entropy numbers, which in turn require good exponential maximal inequalities. Only β -mixing, via decoupling, allows for such an estimate. This is the reason why one can find in the literature the treatment of $\mathbb{Z}_n(g)$ only for β -mixing sequences, see Arcones and Yu (1994) and Doukhan et al. (1995).

The situation for bootstrap of stationary empirical processes is even worse. Although introduced more than twenty years ago (Künsch (1989) and Liu and Singh (1992)), we only found three papers that study the bootstrap for stationary empirical processes indexed by general classes \mathcal{G} , (see also Sengupta et al 2016). All these results operate under β -mixing assumptions. Moreover, only in the case of VC-classes do we have the sharp conditions, see Radulovic (1996). Bracketing classes were considered by Bühlmann (1995), but this was established under very restrictive assumptions on β -mixing coefficients as well as bracketing numbers. It is worth mentioning that the covariance function of the limiting Gaussian process is unknown, and consequently in most actual applications of these results, we heavily rely on the bootstrap version of the process for which adequate results are sorely lacking. In short, the most general α -mixing sequences have never been considered for stationary bootstrap processes $\mathbb{Z}_n^*(g)$, while for the non-bootstrap version $\mathbb{Z}_n(g)$ we have only one example in the literature.

In what follows we prove two general results that allow us to extend the weak limit of classical process $\{\mathbb{G}_n(t), t \in \mathbb{R}\}$ to the weak convergence of $\{\mathbb{Z}_n(g), g \in \mathcal{G}\}$, where \mathcal{G} is a class of functions of uniformly bounded total variation (here called BV_T). We would like to point out that although there are examples of Donsker classes of infinite variation (for instance, $f : [0, 1] \rightarrow [0, 1]$ with $|f(x) - f(y)| \leq |x - y|^\alpha$, $1/2 < \alpha < 1$), such cases are rather the exception than the norm. The majority of examples of bounded Donsker classes that are given in the literature are subsets of the class BV_T .

This enlargement from $\{\mathbb{G}_n(t), t \in \mathbb{R}\}$ to $\{\mathbb{Z}_n(g), g \in \mathcal{G}\}$ is not contingent on the dependence structure of underlying sequence $\{X_i\}$, (or $\{X_i^*\}$) and the only requirement is that $\mathbb{G}_n(t)$ converges weakly to a Gaussian process. This allows us to derive weak convergence of $\{\mathbb{Z}_n(g), g \in \mathcal{G}\}$ and $\{\mathbb{Z}_n^*(g), g \in \mathcal{G}\}$, for α -mixing sequences. The same extension applies for the short memory casual processes considered in Doukhan and Surgailis (1998), as well as processes treated in Dehling (2009). The technique we employ is a simple application of the integration-by-parts formula and the continuous mapping theorem. Arguably this approach has been known before but we could not find this particular application published in general literature, and certainly not among the research related to stationary empirical processes. A follow-up paper (Radulovic,

Wegkamp and Zhao (2016)) extends the results here to classes indexed by multivariate functions of bounded variation, with an emphasis on empirical copula processes. An important technical difference with the follow-up paper is that here we allow for general stationary distribution functions. The proof for general, non-continuous processes, is not trivial caused by technical complications related to the interplay between the atoms of the limiting process $\mathbb{Z}(t)$ and discontinuities of $g \in \mathcal{G}$.

The paper is organized as follows. Section 2 contains the statement of the main result (Theorem 1), and related discussions, while Section 3 contains the statement of Theorem 6, which is a bootstrap version of the same result. Section 4 contains the proofs. For completeness, the well-known integration by parts formula can be found in the appendix.

2 Main results

We use the notation

$$\|g\|_{TV} = \sup_{\Pi} \sum_{x_i \in \Pi} |g(x_i) - g(x_{i-1})|$$

for the total variation norm of a function $g : \mathbb{R} \rightarrow \mathbb{R}$. Here the supremum is taken over all countable partitions $\Pi = \{x_1 < x_2 < \dots\}$ of \mathbb{R} . We set

$$BV_T := \{g : \mathbb{R} \rightarrow \mathbb{R} : \|g\|_{TV} \leq T, \|g\|_{\infty} \leq T\}$$

for $T > 0$. We let $BV'_T \subset BV_T$ be the class of all right-continuous functions $g : \mathbb{R} \rightarrow \mathbb{R}$ with $\|g\|_{TV} \leq T$. Finally we let $\mathbb{Z}_n(t)$, $t \in \mathbb{R}$ be an arbitrary stochastic process such that

A1: $\lim_{|t| \rightarrow \infty} \mathbb{Z}_n(t) = 0$

A2: *The sample paths of $\mathbb{Z}_n(t)$ are right-continuous and of bounded variation.*

Clearly, both requirements A1 and A2 are met for the canonical empirical process $\mathbb{G}_n(t)$. In this paper, we study the limit distribution of the process

$$\bar{\mathbb{Z}}_n(g) := \int g(x) d\mathbb{Z}_n(x), \quad g \in \mathcal{G}$$

for some class $\mathcal{G} \subseteq BV'_T$, for some finite T .

Although the motivation as well as the most notable applications of our results are related to the canonical case $\mathbb{Z}_n(x) = \mathbb{G}_n(x)$, in which case $\bar{\mathbb{Z}}_n(g) = \mathbb{Z}_n(g)$, the actual proof carries over for any process \mathbb{Z}_n as long as the assumptions A1 and A2 are satisfied. Thus, in what follows $\bar{\mathbb{Z}}_n(g) = \int g d\mathbb{Z}_n(x)$ while $\mathbb{Z}_n(g) = \int g d\mathbb{G}_n(x)$.

Our main result is the following theorem.

Theorem 1 *Assume that the stochastic process $\{\mathbb{Z}_n(t), t \in \mathbb{R}\}$ converges weakly to a Gaussian process $\{\mathbb{Z}(t), t \in \mathbb{R}\}$, that is continuous with respect to the distance $d(s, t) = |F_0(s) - F_0(t)|$ for some c.d.f. F_0 . Then, for any $T < \infty$ and $\mathcal{G} \subseteq BV'_T$, the process $\{\bar{\mathbb{Z}}_n(g), g \in \mathcal{G}\}$ converges weakly to a $L_1(F_0)$ -continuous Gaussian process in $\ell^\infty(\mathcal{G})$.*

Remark. The assumption $\|g\|_\infty \leq T$. Namely, $\|g\|_{TV} \leq T$, coupled with the standard envelope assumption $\sup(|g(x)| \leq F(x) \in L^2(P)$, implies that the functions g are uniformly bounded.

Theorem 1 allows us to argue weak convergence of $\mathbb{Z}_n(g)$ via $\mathbb{Z}_n(t)$, regardless of the structure of the latter process. For instance, taking \mathbb{Z}_n as the standard empirical processes \mathbb{G}_n based on stationary sequences X_i , we obtain the following corollary as an immediate consequence of Theorem 1.

Corollary 2 *Let X_k be a stationary sequence of random variables with distribution F and α -mixing coefficients α_n satisfying $\alpha_n = O(n^{-r})$, $n \geq 1$, for some $r > 1$. Then, for any $\mathcal{G} \subseteq BV'_T$, the empirical process $\{\int g d\mathbb{G}_n, g \in \mathcal{G}\}$ converges weakly to a $L_1(F)$ -continuous Gaussian process in $\ell^\infty(\mathcal{G})$. If F is continuous, then \mathcal{G} can be enlarged to BV_T .*

Proof. It is well known, see Theorem 7.2, page 96 in Rio (2000), that $\alpha_n = O(n^{-r})$ with $r > 1$, implies that the standard empirical process $\mathbb{G}_n(t)$ converges weakly to a Brownian bridge process with continuous paths with respect to the distance $d(s, t) = |F(s) - F(t)|$, for the stationary distribution F of X_k . The results for $\mathcal{G} \subseteq BV'_T$ now follows trivially from Theorem 1. If F is continuous, then the limiting Brownian bridge has continuous sample paths with respect to Lebesgue measure on \mathbb{R} , and Lemma 3 below coupled with the continuous mapping theorem now implies the Corollary 2. ■

We used the following lemma.

Lemma 3 *We have*

$$\sup_{g \in BV_T} \inf_{h \in BV'_T} |\bar{\mathbb{Z}}_n(g) - \bar{\mathbb{Z}}_n(h)| \leq T \sup_x |\mathbb{Z}_n(x) - \mathbb{Z}_n(x^-)|.$$

Proof. Let g be an arbitrary function in BV_T . We denote its countable many discontinuities by a_i . Let \bar{g} be the right-continuous version of g , that is, $\bar{g}(x) = g(x)$ for all $x \neq a_i$ and $\bar{g}(a_i) = g(a_i^+)$ for all i . Then

$$\begin{aligned} \left| \int g d\mathbb{Z}_n - \int \bar{g} d\mathbb{Z}_n \right| &\leq \sum_i |g(a_i) - \bar{g}(a_i)| |\mathbb{Z}_n(a_i) - \mathbb{Z}_n(a_i^-)| \\ &\leq \|g\|_{TV} \sup_x |\mathbb{Z}_n(x) - \mathbb{Z}_n(x^-)| \end{aligned}$$

The conclusion now follows easily. ■

We would like to point out that α -mixing is the least restrictive form of available mixing assumptions in the literature. To the best of our knowledge, there are actually very few results that treat processes $\{\mathbb{Z}_n = \int g d\mathbb{G}_n, g \in \mathcal{G}\}$, indexed by functions, and they all require very stringent conditions on the entropy numbers of \mathcal{G} and on the rate of decay for α_k . See, for instance, Andrews and Pollard (1994). This is due to fact that α -mixing does not allow for sharp exponential inequalities for partial sums. Consequently, the only known cases for which we have sharp conditions are under more restrictive, β -mixing dependence. Indeed, β -mixing allows for decoupling and it does yield exponential inequalities not unlike the i.i.d. case. The current state-of-the art results, Arcones & Yu (1994), Doukhan, Massart & Rio (1995), applied to bounded

sequences, require $\sum_n \beta_n < \infty$.

However, Theorem 1 goes beyond dependence defined via mixing conditions. For example, it applies to short memory casual linear sequences. These sequences are defined by

$$X_i = \sum_{j=0}^{\infty} a_j \xi_{i-j}$$

based on i.i.d. random variables ξ_i and constants a_i . While the X_i form a stationary sequence, they do not necessarily satisfy any mixing condition. Weak convergence of the empirical processes $\{\mathbb{G}_n(t), t \in \mathbb{R}\}$ was established under sharp conditions (Doukhan and Surgailis, 1998). To the best of our knowledge, there are no extensions to the more general processes $\mathbb{Z}_n(g)$. Theorem 1 and the Doukhan and Surgailis (1998) result combined imply the following:

Corollary 4 *Let $X_i = \sum_{j \geq 0} a_j \xi_{i-j}$ be such that conditions of Doukhan and Surgailis (1998, pp 87–88) are satisfied and let F be the stationary distribution of X_i . Then, for any $\mathcal{G} \subseteq BV_T$, the empirical process $\{\mathbb{Z}_n = \int g d\mathbb{G}_n, g \in \mathcal{G}\}$ converges weakly to a $L_1(F)$ -continuous Gaussian process in $\ell^\infty(\mathcal{G})$.*

Proof. Doukhan and Surgailis (1998) proved that $\mathbb{G}_n(t)$, converges weakly to a Gaussian process in the Skorohod's space. Since F is continuous under their assumptions, see Doukhan and Surgailis (1998, pp 88) the limiting process of \mathbb{G}_n is continuous with respect to the distance $d(s, t) = |F(s) - F(t)|$ for the stationary distribution F . The proof now follows trivially from Theorem 1 and Lemma 3. ■

The recent papers by Dehling et al. (2009 and 2014) offer yet another, clever way to prove the weak limit of the standard empirical processes \mathbb{G}_n based on stationary sequences that are not necessarily mixing. Their technique uses finite dimensional convergence coupled with a bound on the higher moments of partial sums, which in turn controls the dependence structure. Dehling et al. (2009) establishes the weak convergence of \mathbb{G}_n , while Dehling et al. (2014) extends this idea to more general classes of functions. However, the authors impose cumbersome entropy conditions and

only manage to marginally extend the classes. For example, they manage to prove weak convergence of the process $\int f_t d\mathbb{G}_n$, indexed by functions $f_t(x)$ which constitute one-dimensional monotone class (with restrictive requirement that $s \leq t \Rightarrow f_s \leq f_t$). Theorem 1 applied in their setting, yields a more general result.

Corollary 5 *Let $\mathcal{G} \subseteq BV_T$ and let F be the stationary distribution of X_i . Under assumptions (i) and (ii) in Section 1 of Dehling et al. (2009), the empirical process $\{\int g d\mathbb{G}_n, g \in \mathcal{G}\}$ converges weakly to a $L_1(F)$ -continuous Gaussian process in $\ell^\infty(\mathcal{G})$.*

Proof. The underlying distribution function F of X_k in Dehling et al (2009) is continuous, see display (3) at page 3702. Moreover they established the weak convergency of the process \mathbb{G}_n to a Gaussian process that is continuous with respect to distance $d(s, t) = |F(s) - F(t)|$. The proof follows trivially from Theorem 1 and Lemma 3. ■

3 Bootstrap

The limit of $\mathbb{Z}_n(g)$ in Theorem 1 is a Gaussian process with complicated covariance structure

$$\begin{aligned} \mathbb{E}[\mathbb{Z}(f)\mathbb{Z}(g)] &= \text{Cov}(f(X_0), g(X_0)) + \sum_{k=1}^{\infty} \text{Cov}(f(X_0), g(X_k)) \\ &\quad + \sum_{k=1}^{\infty} \text{Cov}(f(X_k), g(X_0)). \end{aligned}$$

Consequently, the actual applications of the weak limit results of $\mathbb{Z}_n(g)$ are very hard to implement and very seldom a closed form solution is available. This situation calls for the bootstrap principle. That is, given the sample X_1, \dots, X_n , we let \mathbb{G}_n^* be bootstrap empirical process

$$\mathbb{G}_n^*(t) = \sqrt{m_n} \left(\frac{1}{m_n} \sum_{i=1}^{m_n} 1_{(-\infty, t]}(X_i^*) - \frac{1}{n} \sum_{i=1}^n 1_{(-\infty, t]}(X_i) \right), \quad t \in \mathbb{R},$$

based on a bootstrap sample $X_1^*, \dots, X_{m_n}^*$. We stress that for our purposes here no additional assumption on the structure of the variables $X_{i,n}^*$ is required. Analogous to

$\mathbb{Z}_n(g) = \int g d\mathbb{G}_n$, we define $\mathbb{Z}_n^*(g) := \int g d\mathbb{G}_n^*$ for any $g \in BV_T'$ with $T < \infty$. Recall that the bounded Lipschitz distance

$$d_{BL}(\mathbb{Z}_n^*, \mathbb{Z}) = \sup_{h \in BL_1} |\mathbb{E}^*[h(\mathbb{Z}_n^*)] - \mathbb{E}[h(\mathbb{Z})]|$$

between two processes \mathbb{Z}_n^* and \mathbb{Z} metrizes weak convergence. Here \mathbb{E}^* is the expectation over the randomness of the bootstrap sample $X_1^*, \dots, X_{m_n}^*$, conditionally given the original sample X_1, \dots, X_n , and BL_1 is the space of Lipschitz functionals $h : \ell^\infty(\mathcal{G}) \rightarrow \mathbb{R}$ with $|h(x)| \leq 1$ and $|h(x) - h(y)| \leq \|x - y\|_\infty$ for all $x, y \in \mathcal{G} \subset BV_T$. Customary in the literature, if the random variable $d_{BL}(\mathbb{Z}_n^*, \mathbb{Z})$ converges to zero in probability, we speak of weak convergence in probability; if it converges to zero almost surely, we speak of weak convergence almost surely.

Theorem 6 *Let $\mathcal{G} \subseteq BV_T'$. Assume that, conditionally on X_1, \dots, X_n , in probability, $\{G_n^*(t), t \in R\}$ converges weakly to a Gaussian process that is continuous with respect to the distance $d(s, t) = |F(s) - F(t)|$ based on the stationary distribution F of X_i . Then, the process $\mathbb{Z}_n^* = \{\int g d\mathbb{G}_n^*, g \in \mathcal{G}\}$ converges weakly to a $L_1(F)$ -continuous Gaussian process in $\ell^\infty(\mathcal{G})$.*

If the weak convergence of \mathbb{G}_n^ holds almost surely, then the conclusion that \mathbb{Z}_n^* converges weakly in $\ell^\infty(\mathcal{G})$ holds almost surely as well.*

Moreover, if \mathbb{G}_n and \mathbb{G}_n^ converge to the same limit, then \mathbb{Z}_n and \mathbb{Z}_n^* converge to the same limit in $\ell^\infty(\mathcal{G})$.*

The literature offers numerous bootstrapping techniques for stationary data: moving block bootstrap, stationary bootstrap, sieved bootstrap, Markov chain bootstrap, etc., but, their validity is proved for specific cases/statistics only. Due to complications with entropy calculations for dependent triangular arrays, almost all results treat the standard empirical processes \mathbb{G}_n^* with few notable exceptions. The moving block bootstrap was justified for VC-type classes, but only under rather restrictive beta-mixing conditions on X_i (Radulović, 1996). Bracketing classes were considered by Bühlmann (1995), but his conditions are even more restrictive. In contrast, the process $\{\mathbb{G}_n(t), t \in \mathbb{R}\}$ is rather easy to bootstrap. This coupled with Theorem 6 offers the following result.

Corollary 7 *Let X_j be a stationary sequence of random variables with continuous stationary c.d.f. F and α -mixing coefficients satisfying $\sum_{k \geq n} \alpha_k = O(n^{-\gamma})$, for some $0 < \gamma < 1/3$. Let GG_n^* be the bootstrapped standard empirical process based on the moving block bootstrap, with block sizes b_n , specified in Peligrad (1998, p 882). Then, for $\mathcal{G} \subset BV_T$, the bootstrap empirical process $\mathbb{Z}_n^* = \{\int g d\mathbb{G}_n^*, g \in \mathcal{G}\}$ converges weakly to a $L_1(F)$ -continuous Gaussian process, almost surely.*

Proof. Theorem 2.3 of Peligrad (1998) establishes the convergence of \mathbb{G}_n^* to a Gaussian process that is continuous with respect to the distance $d(s, t) = |F(s) - F(t)|$ based on the stationary distribution F of X_i . Invoke Theorem 6 and Lemma 3 to conclude the proof. ■

Just as for weak convergence of the empirical process based on stationary sequences, there are numerous results that consider bootstrap for stationary, non-mixing sequences. For example, Ktaibi et al. (2014) study short memory casual linear sequences, and prove weak convergence of \mathbb{G}_n^* under conditions akin to the ones required for its non-bootstrap counterpart \mathbb{G}_n (Doukhan and Surgailis, 1998). Again, Theorem 6 easily extends this result.

Corollary 8 *Let $X_i = \sum_{j \geq 0} a_j \xi_{i-j}$ be a sequence of random variables with stationary distribution F such that conditions of Ktaibi et al (2014) are satisfied. Then, for any $\mathcal{G} \subset BV_T$, the process $\mathbb{Z}_n^* = \{\int g d\mathbb{G}_n^*, g \in \mathcal{G}\}$ converges weakly a.s. to a $L_1(F)$ -continuous Gaussian limit.*

Proof. Ktaibi et al (2014) proved that the convergence of \mathbb{G}_n^* in the Skorohod's space, for continuous F , which in turn implies that the limiting process is continuous with respect to distance $d(s, t) = |F(s) - F(t)|$. Again, invoke Theorem 6 and Lemma 3 to conclude the proof. ■

4 Proofs for Theorems 1 & 6

The proofs of Theorems 1 and 6 follow easily if the limit $\mathbb{Z}(t)$ of $\mathbb{Z}_n(t)$ is continuous with respect to Lebesgue measure. Namely, we let d_{BL} be the bounded Lipschitz

metric that metrizes the weak convergence, see, e.g., Van der Vaart & Wellner (1996, page 73) for the definition. Hence we need to show $d_{BL}(\mathbb{Z}_n, \mathbb{Z}) \rightarrow 0$ as $n \rightarrow \infty$. Set $\bar{\mathbb{Z}}_n(g) = \int g d\mathbb{Z}_n$ for any $g \in BV'_T$. The assumptions A1 and A2 imply that the Lebesgue Stieltjes integrals

$$\tilde{\mathbb{Z}}_n(g) = - \int \mathbb{Z}_n dg$$

are well defined. Next, by the integration by parts formula in Lemma A, we have

$$\bar{\mathbb{Z}}_n(g) = \tilde{\mathbb{Z}}_n(g) + R_n(g)$$

with

$$R_n(g) \leq T \sup_x |\mathbb{Z}_n(x) - \mathbb{Z}_n(x^-)| \rightarrow 0,$$

in probability, as $n \rightarrow \infty$, and \mathbb{Z} is continuous. For fixed g , $\tilde{\mathbb{Z}}_n(g)$ converges weakly to $\tilde{\mathbb{Z}}(g) := - \int \mathbb{Z} dg$ by the continuous mapping theorem and weak convergence of \mathbb{Z}_n . The continuous mapping theorem also guarantees that the limit $\tilde{\mathbb{Z}}$ is tight in $\ell^\infty(\mathcal{G})$ as the map $\Gamma : \ell^\infty(\mathbb{R}) \rightarrow \ell^\infty(\mathcal{G})$ defined as $\Gamma_f(X) = - \int X dg$, $g \in \mathcal{G}$, is continuous. By the triangle inequality,

$$\begin{aligned} d_{BL}(\bar{\mathbb{Z}}_n, \tilde{\mathbb{Z}}) &\leq d_{BL}(\bar{\mathbb{Z}}_n, \tilde{\mathbb{Z}}_n) + d_{BL}(\tilde{\mathbb{Z}}_n, \tilde{\mathbb{Z}}) \\ &= \sup_{H \in BL_1} |\mathbb{E}H(\bar{\mathbb{Z}}_n) - \mathbb{E}H(\tilde{\mathbb{Z}}_n)| + \sup_{H \in BL_1} |\mathbb{E}H(\tilde{\mathbb{Z}}_n) - \mathbb{E}H(\tilde{\mathbb{Z}})| \\ &\leq T\mathbb{E}[\sup_x |\mathbb{Z}_n(x) - \mathbb{Z}_n(x^-)|] + T \sup_{H' \in BL_1} |\mathbb{E}H'(\tilde{\mathbb{Z}}_n) - \mathbb{E}H'(\tilde{\mathbb{Z}})| \\ &= T\mathbb{E}[\sup_x |\mathbb{Z}_n(x) - \mathbb{Z}_n(x^-)|] + Td_{BL}(\mathbb{Z}_n, \mathbb{Z}) \end{aligned}$$

The second term follows since the map $\Gamma_f(X) := \int X df$ is Lipschitz with Lipschitz constant $\int |df| \leq T$ and the suprema are taken over all Lipschitz functionals $H : \ell^\infty(\mathcal{G}) \rightarrow \mathbb{R}$ with $\|H\|_\infty \leq 1$ and $|H(X) - H(Y)| \leq \|X - Y\|_\infty$ and $H' : \ell^\infty(\mathbb{R}) \rightarrow \mathbb{R}$ with $\|H'\|_\infty \leq 1$ and $|H'(X) - H'(Y)| \leq \|X - Y\|_\infty$, respectively. Together with the tightness of the limit $\tilde{\mathbb{Z}}$, this implies that the empirical process $\bar{\mathbb{Z}}_n(g)$ indexed by $g \in \mathcal{G} \subset BV'_T$ converges weakly.

The above proof for continuous limit processes \mathbb{Z} is rather simple. Nevertheless, we could not find an actual publication of this trick. We would like to stress that the

extension of this proof to general, non-continuous processes, is not trivial. Technical complications related to the interplay between the atoms of the limiting process $\mathbb{Z}(t)$ and discontinuities of $g \in \mathcal{G}$, require some care.

Lemma 9 .Assume that the stochastic process $\mathbb{Z}_n(t)$ converges weakly to a Gaussian process $\mathbb{Z}(t)$. Then, for any right continuous function $g : \mathbb{R} \rightarrow \mathbb{R}$ of bounded variation $\overline{\mathbb{Z}}_n(g) = \int g d\mathbb{Z}_n$ is well defined, and converges to a normal distribution on \mathbb{R} .

Proof. Let g be an arbitrary right-continuous function of bounded variation. First, we notice that $\int g d\mathbb{Z}_n$ and $\int \mathbb{Z}_n dg$ are indeed well defined as Lebesgue-Stieltjes integrals. Recall that g can have only countably many discontinuities which we will denote a_i . By the integration by parts formula Lemma A in appendix, we have

$$\int g(x) d\mathbb{Z}_n(x) = T_1(\mathbb{Z}_n) + T_2(\mathbb{Z}_n)$$

with operators $T_1, T_2 : \ell^\infty(\mathbb{R}) \rightarrow \mathbb{R}$

$$\begin{aligned} T_1(\mathbb{Z}_n) &:= - \int \mathbb{Z}_n(x) dg(x) \\ T_2(\mathbb{Z}_n) &:= \int \int 1_{x=y} dg(x) d\mathbb{Z}_n(y) = \sum_i (g(a_i) - g(a_i^-)) (\mathbb{Z}_n(a_i) - \mathbb{Z}_n(a_i^-)). \end{aligned}$$

Since g has finite variation, it is bounded with $\sum_i |\alpha_i| < \infty$ and $\sum_i \alpha_i^2 < \infty$ for the jumps $\alpha_i := g(a_i) - g(a_i^-)$. Hence, the operator

$$T_2(\mathbb{Z}_n) = \sum_i \alpha_i (\mathbb{Z}_n(a_i) - \mathbb{Z}_n(a_i^-))$$

is linear. We conclude the proof by observing that the linearity of the operators T_1 and T_2 and the weak convergence of $\mathbb{Z}_n(t)$ to a Gaussian process \mathbb{Z} ensures that these sequences $Y_n := \int g d\mathbb{Z}_n = T_1(\mathbb{Z}_n) + T_2(\mathbb{Z}_n)$ converge weakly to normal distribution via the continuous mapping theorem. ■

For any c.d.f. F_0 and any $\beta > 0$, we define

$$\begin{aligned} \Psi_{\beta, F_0}(\mathbb{Z}_n) &:= \sup_{|F_0(s) - F_0(t)| \leq \beta} |\mathbb{Z}_n(s) - \mathbb{Z}_n(t)|, \\ \tilde{\mathfrak{O}}_{\beta, F_0}(\mathbb{Z}_n) &:= \sup_{F_0(s) - F_0(s^-) > \beta} |\mathbb{Z}_n(s) - \mathbb{Z}_n(s^-)|, \end{aligned}$$

with the convention that sup over empty set is equal zero.

Clearly, $\bar{\mathfrak{O}}_{\beta, F_0}(\mathbb{Z}_n) = 0$ for all $\beta > 0$ if F_0 is continuous. In general, for arbitrary F_0 , the quantity $\bar{\mathfrak{O}}_{\beta, F_0}(\mathbb{Z}_n)$ is bounded in probability, for all $\beta > 0$, by the continuous mapping theorem, as long as \mathbb{Z}_n converges weakly. The following lemma plays an instrumental role and it could be of independent interest.

Lemma 10 (Decoupling Lemma) *For any distribution function F_0 on \mathbb{R} , any right-continuous function $g : \mathbb{R} \rightarrow \mathbb{R}$, and any $\beta > 0$, we have*

$$|\bar{\mathbb{Z}}_n(g)| \leq \{2\beta^{-1}\|g\|_{L_1(F_0)} + 6\|g\|_{TV}\} \Psi_{2\beta, F_0}(\mathbb{Z}_n) + \beta^{-1}\|g\|_{L_1(F_0)} \bar{\mathfrak{O}}_{\beta, F_0}(\mathbb{Z}_n).$$

Here $\|g\|_{L_1(F_0)} = \int |g| dF_0$ and \mathbb{Z}_n satisfies the assumptions A1 and A2.

Proof. Without loss of generality we can assume that $\|g\|_{L_1(F_0)} + \|g\|_{TV} < \infty$. Since F_0 is a distribution function we can construct, for any $0 < \beta < 1$, a finite grid $-\infty = s_0 < s_1 < \dots < s_M < \infty$ such that

$$F_0(s_j) - F_0(s_{j-1}) \geq \beta, \quad F_0(s_M) < 1 - 2\beta$$

and

$$F_0(s_j^-) - F_0(s_{j-1}) \leq 2\beta,$$

leaving the possible jumps $F_0(s_j) - F_0(s_j^-)$ unspecified. Based on this grid we approximate $\mathbb{Z}_n(t)$ by

$$\tilde{\mathbb{Z}}_n(t) := \sum_{i=1}^M \mathbb{Z}_n(s_{j-1}) 1_{[s_{j-1}, s_j)}(t)$$

and we set $\tilde{\mathbb{Z}}_n(\pm\infty) = 0$. We observe that by construction

$$\begin{aligned} \sup_x |\mathbb{Z}_n(x) - \tilde{\mathbb{Z}}_n(x)| &\leq \max_{1 \leq j \leq M} \sup_{x \in [s_{j-1}, s_j)} |\mathbb{Z}_n(x) - \mathbb{Z}_n(s_{j-1})| + \sup_{x \geq s_M} |\mathbb{Z}_n(x)| \\ &\leq \sup_{|F_0(x) - F_0(y)| \leq 2\beta} |\mathbb{Z}_n(x) - \mathbb{Z}_n(y)| = \Psi_{\beta, F_0}(\mathbb{Z}_n) \end{aligned} \quad (3)$$

Since the process $\tilde{\mathbb{Z}}_n$ inherits the bounded variation properties of \mathbb{Z}_n ,

$$\int g(s) d\mathbb{Z}_n(s) = \int g(s) d\tilde{\mathbb{Z}}_n(s) + \int g(s) d(\mathbb{Z}_n - \tilde{\mathbb{Z}}_n)(s).$$

is well defined for any right-continuous function g of bounded variation. Using the integration by parts formula Lemma A we obtain for the last term on the right

$$\int g(s)d(\mathbb{Z}_n - \tilde{\mathbb{Z}}_n)(s) = - \int (\mathbb{Z}_n - \tilde{\mathbb{Z}}_n)(s)dg(s) + \int \int 1_{x=y}dg(x)d(\mathbb{Z}_n - \tilde{\mathbb{Z}}_n)(y).$$

Since g is of bounded variation, it has countably many discontinuities a_i . Using (3), we obtain

$$\begin{aligned} \left| \int \int 1_{x=y}dg(x)d(\mathbb{Z}_n - \tilde{\mathbb{Z}}_n)(y) \right| &\leq \sum_i |g(a_i) - g(a_i^-)| \left| (\mathbb{Z}_n - \tilde{\mathbb{Z}}_n)(a_i) - (\mathbb{Z}_n - \tilde{\mathbb{Z}}_n)(a_i^-) \right| \\ &\leq 2 \|g\|_{TV} \Psi_{2\beta, F_0}(\mathbb{Z}_n). \end{aligned}$$

Consequently,

$$\left| \int g(s)d(\mathbb{Z}_n - \tilde{\mathbb{Z}}_n)(s) \right| \leq 3 \|g\|_{TV} \Psi_{2\beta, F_0}(\mathbb{Z}_n)$$

Next we deal with the finite dimensional approximation

$$\int g(s)d\tilde{\mathbb{Z}}_n(s) = \sum_{j=1}^M g(s_j)(\mathbb{Z}_n(s_j) - \mathbb{Z}_n(s_{j-1})).$$

Clearly,

$$\left| \int g(s)d\tilde{\mathbb{Z}}_n(s) \right| \leq \left| \sum_{j=1}^M g(s_j)(\mathbb{Z}_n(s_j^-) - \mathbb{Z}_n(s_{j-1})) \right| + \left| \sum_{j=1}^M g(s_j)(\mathbb{Z}_n(s_j) - \mathbb{Z}_n(s_j^-)) \right| \quad (4)$$

For the first term in (4) we introduce the step function

$$g^*(t) = \sum_{j=1}^M 1_{(s_{j-1}, s_j]}(t) \inf_{s_{j-1} < s \leq s_j} |g(s)|.$$

designed to approximate $|g(t)|$. Clearly $0 \leq g^*(t) \leq |g(t)|$ and

$$\begin{aligned} \sum_{j=1}^M g^*(s_j) &\leq \beta^{-1} \sum_{j=1}^M g^*(s_j)(F_0(s_j) - F_0(s_{j-1})) \\ &= \beta^{-1} \int g^*dF_0 \\ &\leq \beta^{-1} \int |g|dF_0 = \beta^{-1} \|g\|_{L_1(F_0)} \end{aligned}$$

Hence, by (3) we have

$$\begin{aligned} \left| \sum_{j=1}^M g(s_j) (\mathbb{Z}_n(s_j^-) - \mathbb{Z}_n(s_{j-1})) \right| &\leq \Psi_{2\beta, F_0}(\mathbb{Z}_n) \left[\sum_{j=1}^M (|g(s_j)| - g^*(s_j)) + \sum_{j=1}^M g^*(s_j) \right] \\ &\leq \Psi_{2\beta, F_0}(\mathbb{Z}_n) (\|g\|_{TV} + \beta^{-1} \|g\|_{L_1(F_0)}) \end{aligned} \quad (5)$$

For the second term in (4), we have

$$\begin{aligned} &\left| \sum_{j=1}^M g(s_j) (\mathbb{Z}_n(s_j)) - \mathbb{Z}_n(s_j^-) \right| \\ &\leq \left| \sum_{j: F_0(s_j) - F_0(s_j^-) \leq \beta} g(s_j) (\mathbb{Z}_n(s_j)) - \mathbb{Z}_n(s_j^-) \right| + \left| \sum_{j: F_0(s_j) - F_0(s_j^-) > \beta} g(s_j) (\mathbb{Z}_n(s_j)) - \mathbb{Z}_n(s_j^-) \right| \\ &\leq \Psi_{2\beta, F_0}(\mathbb{Z}_n) (2\|g\|_{TV} + \beta^{-1} \|g\|_{L_1(F_0)}) + \check{\mathfrak{D}}_{\beta, F_0}(\mathbb{Z}_n) \beta^{-1} \|g\|_{L_1(F_0)}, \end{aligned} \quad (6)$$

using for the last inequality

$$\sum_{j=1}^M |g(s_j)| \leq 2\|g\|_{TV} + \beta^{-1} \|g\|_{L_1(F_0)}$$

and

$$\begin{aligned} \sum_{j=1}^M |g(s_j)| 1\{F_0(s_j) - F_0(s_j^-) > \beta\} &\leq \beta^{-1} \sum_{j=1}^M |g(s_j)| (F_0(s_j) - F_0(s_j^-)) \\ &\leq \beta^{-1} \|g\|_{L_1(F_0)}. \end{aligned}$$

The proof for Lemma 10 now follows by combining the estimates (3) and (4) and (5). \blacksquare

An immediate corollary is the following result.

Corollary 11 *For any c.d.f. F_0 , and for all $T < \infty$, $\delta > 0$ and $p \geq 1$, we have*

$$\sup_{\|g\|_{L_p(F_0)} \leq \delta, \|g\|_{TV} \leq T} \left| \int g d\mathbb{Z}_n \right| \leq (2\sqrt{\delta} + 6T) \Psi_{2\sqrt{\delta}, F_0}(\mathbb{Z}_n) + \check{\mathfrak{D}}_{\sqrt{\delta}, F_0}(\mathbb{Z}_n) \sqrt{\delta},$$

where $\|g\|_{L_p(F_0)} = (\int |g|^p dF_0)^{1/p}$ and the sup is taken over all right-continuous functions g .

Proof. The proof follows trivially from Lemma 10 by taking $\beta = \sqrt{\delta}$ and observing that $\|g\|_{L_1(F_0)} \leq \|g\|_{L_p(F_0)}$ for $p \geq 1$. \blacksquare

Proof of Theorem 1.

First we recall that, (see for instance, Van der Vaart & Wellner; 1996, Chapter 1.5), the process $\{\bar{\mathbb{Z}}_n(g), g \in \mathcal{G}\}$ converges weakly to a tight limit in $\ell^\infty(\mathcal{G})$, provided

- (a) the marginals $(\bar{\mathbb{Z}}_n(g_1), \dots, \bar{\mathbb{Z}}_n(g_k))$ converge weakly for every finite subset $g_1, \dots, g_k \in \mathcal{G}$, and
- (b) there exists a semi-metric ρ on \mathcal{G} such that (\mathcal{G}, ρ) is totally bounded and $\bar{\mathbb{Z}}_n(g)$ is ρ -stochastically equicontinuous, that is,

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P} \left\{ \sup_{\rho(f,g) \leq \delta} |\bar{\mathbb{Z}}_n(f) - \bar{\mathbb{Z}}_n(g)| > \varepsilon \right\} = 0$$

for all $\varepsilon > 0$.

The finite dimensional convergence (a), follows trivially from Lemma 9, linearity of the process $\bar{\mathbb{Z}}_n(g)$ and Cramer-Wold device. As for stochastic equicontinuity (b) of $\bar{\mathbb{Z}}_n(g)$, it is sufficient to show that, for every $\varepsilon > 0$,

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P} \left\{ \sup_{\|h\|_{L_1(F_0)} \leq \delta} \left| \int h(t) d\mathbb{Z}_n(t) \right| > \varepsilon \right\} = 0,$$

here the supremum is taken over all differences $h = g - g'$ with $g, g' \in \mathcal{G}$ and $\|h\|_{L_1(F_0)} = \int |h| dF_0 \leq \delta$. Since h is also right continuous, an $\|h\|_{TV} \leq 2T$. Corollary 11 implies that

$$\sup_{\|h\|_{L_1(F_0)} \leq \delta} \left| \int h(t) d\mathbb{Z}_n(t) \right| \leq (2\sqrt{\delta} + 12T) \Psi_{2\sqrt{\delta}, F_0}(\mathbb{Z}_n) + \bar{\Psi}_{\sqrt{\delta}, F_0}(\mathbb{Z}_n) \sqrt{\delta}.$$

Let $f_t(x) = 1\{x \leq t\}$, so that $\mathbb{Z}_n(t) = \bar{\mathbb{Z}}_n(f_t)$ and

$$d(s, t) = |F_0(s) - F_0(t)|$$

and observe that

$$\sup_{d(s,t) \leq \delta} |\mathbb{Z}_n(s) - \mathbb{Z}_n(t)| = \Psi_{\delta, F_0}(\mathbb{Z}_n).$$

The weak convergence of $\mathbb{Z}_n(t)$, or equivalently, $\bar{\mathbb{Z}}_n(f_t)$, to a continuous (with respect to $d(.,.)$) process $\mathbb{Z}(t)$ implies that

$$\Psi_{2\sqrt{\delta}, F_0}(\mathbb{Z}_n) \xrightarrow{P} 0 \text{ as } \delta \rightarrow 0 \text{ and } n \rightarrow \infty.$$

Moreover, the weak convergence of $\mathbb{Z}_n(t)$ implies that $\bar{\mathfrak{D}}_{\sqrt{\delta}, F_0}(\mathbb{Z}_n)$ is bounded in probability, so $\bar{\mathfrak{D}}_{\sqrt{\delta}, F_0}(\mathbb{Z}_n)\sqrt{\delta} \rightarrow 0$ as $\delta \rightarrow 0$ and $n \rightarrow \infty$.

Summarizing, the process $\bar{\mathbb{Z}}_n(g)$, converges for each $g \in \mathcal{G}$ to a Gaussian limit and it is uniformly $L_1(F_0)$ –equicontinuous, in probability. Moreover, $(\mathcal{G}, L_1(F_0))$ is totally bounded (for any c.d.f. F_0). This implies the weak convergence of $\bar{\mathbb{Z}}_n(g)$ to a $L_1(F_0)$ –continuous process (see Theorems 1.54 and 1.5.7 Van der Vaart & Wellner 1996).

Proof of Theorem 6.

Here we only prove the “in probability” case. The almost sure statement follows after straightforward changes in the proof of Theorem 1. A simple modification of Lemma 9 yields

$$\int g(t) d\mathbb{G}_n^*(t) = T_1(\mathbb{G}_n^*) + T_2(\mathbb{G}_n^*)$$

for the same operators T_1 and T_2 as defined in Lemma 9. Since \mathbb{G}_n^* converges to a Gaussian limit the finite dimensional convergence follows by recalling the computation presented in the proof of Lemma 9 (by replacing $\mathbb{G}_n(t)$ with $\mathbb{G}_n^*(t)$). As for stochastic equicontinuity of $\mathbb{Z}_n^*(g)$ we find, analogous to Corollary 11, that for $\delta > 0$

$$\sup_{\|g\|_{L_P(F_0)} \leq \delta, \|g\|_{TV} \leq T} \left| \int g d\mathbb{G}_n^* \right| \leq (2\sqrt{\delta} + 12T) \Psi_{2\sqrt{\delta}, F_0}(\mathbb{G}_n^*) + \bar{\mathfrak{D}}_{\sqrt{\delta}, F_0}(\mathbb{G}_n^*)\sqrt{\delta}.$$

Now the weak convergence of \mathbb{Z}_n^* , follows from the convergence of \mathbb{G}_n^* , as in the proof of Theorem 1. Moreover, if $\mathbb{G}_n(t)$ and $\mathbb{G}_n^*(t)$ converge to the same Gaussian process, than Lemma 9 coupled with Cramer-Wold device implies that the finite dimensional distribution of the limiting process of $\mathbb{Z}_n(g)$ and $\mathbb{Z}_n^*(g)$ are the same. This concludes the proof.

A Appendix

For completeness, we state the following classical result and give a simple elementary proof which was communicated to us by David Pollard.

Lemma A. *Let f and g be right-continuous functions of bounded variation and define measures μ and ν as $\mu(-\infty, x] = f(x) - f(-\infty)$ and $\nu(-\infty, y] = g(y) - g(-\infty)$. Then*

$$\int f(x) dg(x) + \int g(x) df(x) = (fg)(\infty) - (fg)(-\infty) + \int \int 1_{x=y} d\mu d\nu.$$

Moreover, if either $f(\pm\infty) = 0$ or $g(\pm\infty) = 0$, then

$$\int f(x) dg(x) + \int g(x) df(x) = \int \int 1_{x=y} d\mu d\nu.$$

Proof. Set $H(x, y) = 1\{x \leq y\}$ and observe that by the very definition of Lebesgue integral

$$f(y) = \int H(x, y) d\mu(x) + f(-\infty)$$

and

$$g(x) = \int H(y, x) d\nu(y) + g(-\infty).$$

Hence

$$\begin{aligned} & \int f(y) d\nu(y) + \int g(x) d\mu(x) \\ &= \int \left(\int H(x, y) d\mu(x) \right) d\nu(y) + \int \left(\int H(y, x) d\nu(y) \right) d\mu(x) \\ & \quad + f(-\infty)(g(\infty) - g(-\infty)) + g(-\infty)(f(\infty) - f(-\infty)) \end{aligned}$$

Next we apply Fubini

$$\begin{aligned} & \int \left(\int H(x, y) d\mu(x) \right) d\nu(y) + \int \left(\int H(y, x) d\nu(y) \right) d\mu(x) \\ &= \int \int (H(x, y) + H(y, x)) d\mu(x) d\nu(y) = \int \int (1_{x \leq y} + 1_{y \leq x}) d\mu(x) d\nu(y) \\ &= \int \int d\mu(x) d\nu(y) + \int \int 1_{x=y} d\mu(x) d\nu(y) \end{aligned}$$

to prove the lemma. ■

References

- [1] D.W.K. Andrews and D.P. Pollard (1994). An Introduction to Functional Central Limit Theorems for Dependent Stochastic Processes. *International Statistical Review*, **62**(1), 119–132.
- [2] M.A. Arcones, and B. Yu (1994). Central limit theorems for empirical and U-processes of stationary mixing sequences. *Journal of Theoretical Probability* **7**, 47–71.
- [3] P. Bühlmann (1995). The blockwise bootstrap for general empirical processes of stationary sequences. *Stochastic Processes and their Applications*, **58**, 247–265.
- [4] R. M. Dudley, (1999), *Uniform Central Limit Theorems*, Cambridge University Press.
- [5] H. Dehling, O. Durieu and D. Volny (2009). New techniques for empirical processes of dependent data. *Stochastic Processes and their Applications*, **119**, 3699–3718.
- [6] H. Dehling, O. Durieu and M. Tusche (2014). Approximating class approach for empirical processes of dependent sequences indexed by functions. *Bernoulli*, **20**, 1372–1403.
- [7] P. Doukhan, and D. Surgailis (1998). Functional central limit theorem for the empirical process of short memory linear processes. *Comptes Rendus de l'Académie des Sciences, Mathématique*, **326**, 87–92.
- [8] F. El Ktaibi, B.G. Ivanoff, N.C. Weber (2014). Bootstrapping the empirical distribution of a linear process. *Statistics & Probability Letters*, **93**, 134–142
- [9] T.H. Hildebrandt (1963). *Introduction to the Theory of Integration*. Academic Press, New York, London.
- [10] R.Y. Liu and K. Singh (1992), Moving blocks jackknife and bootstrap capture weak dependence, in: R. LePage and L. Billard, eds., *Exploring the Limits of Bootstrap*, Wiley, New York, pp. 225–248.

- [11] H.R. Kunsch (1989), The jackknife and the bootstrap for general stationary observations, *Annals of Statistics*, **17**, 1217–1241.
- [12] M. Peligrad (1998). On the blockwise bootstrap for empirical processes for stationary sequences. *Annals of Probability*, **26**, 887–901.
- [13] D. Radulović (1996). The bootstrap for empirical processes based on stationary observations. *Stochastic Processes and Their Applications*, **65**, 259–279.
- [14] D. Radulović, M. Wegkamp and Y. Zhao (2016). Weak convergence of empirical copula processes indexed by functions. *Bernoulli*, in press.
- [15] E. Rio (1998). Processus empiriques absolument régulllets et entropie universelle. *Probability Theory and Related Fields*, **111**, 585–608.
- [16] E. Rio (2000). *Théorie asymptotique des processus aléatoires faiblement dépendants*, Collection Mathématiques and Applications, **31**. Springer.
- [17] S. Sengupta, S. Volgushev & X. Shao (2016), A Subsampled Double Bootstrap for Massive Data, *JASA* **111**, Issue 515
- [18] G.R. Shorack and J.A. Wellner (2009). *Empirical Processes with Applications to Statistics*, 2nd Edition, SIAM.
- [19] A.W. van der Vaart and J. Wellner (2000). *Weak convergence and empirical processes*. Springer Series in Statistics. Springer.
- [20] D. Viennet (1997). Inequalities for absolutely regular sequences: application to density estimation. *Probability Thoery and Related Fields*, **107**, 467–492.